图书情报工作
第 65 卷 第 12 期 2021 年 6 月

知识产权视角下科学数据安全管理的策略选择*

■ 罗娇[1] 刘细文[2,3]
[1] 中国农业大学人文与发展学院 北京 100083 [2] 中国科学院文献情报中心 北京 100190
[3] 中国科学院大学经济管理学院图书情报与档案管理系 北京 100190

摘 要:[目的/意义]科学数据安全是一个复杂、动态的概念,包含了主权安全、产权安全和共享安全,其中产权安全又是主权安全、共享安全的重要前提。在我国未就数据产权立法的背景下,以现行法中与科学数据产权存在交叉关系的知识产权为切入点,探析科学数据安全管理中的产权关系具有重要意义。[方法/过程]针对科学数据产权缺位、主权缺失、共享受阻问题,以法律规范分析为基础,运用法律解释方法,分析知识产权与科学数据的关系,明确产权是解决科学数据安全的关键措施,并结合国内外实践案例探索可能的解决方案。[结果/结论]在产权问题上,以知识产权为基础,合同为辅助,通过权利归属条款、许可授权条款、保密条款和利益分配条款保障科学数据的产权安全。在主权问题上,构建国家层面的科学数据管理顶层设计,增强数据主权意识;建立机构层面的科学数据管理与共享政策,避免数据违规外流。在共享问题上,依据不同类型的知识产权特征来清除科学数据开放共享的权利障碍,通过许可协议实现科学数据的开放共享。

关键词:科学数据 知识产权 合同 数据安全
分类号:G251
DOI:10.13266/j.issn.0252-3116.2021.12.003

数据时代,数据定义万物、链接万物、变革万物。科研活动中亦汇集了大量数据。科学研究范式从描述自然现象的"经验科学"、使用模型或归纳法的"理论科学"、模拟复杂现象的"计算科学",跃升至融合理论、实验和模拟的"数据探索"[1-2],形成数据密集型科学发现研究范式:欧洲大型强子对撞机的数据确定了希格斯玻色子的存在;基因大数据为基因组学研究提供了新的机会;时空大数据在全球环境变化研究中发挥了重大作用[3],"数据驱动科学"成为趋势。科学数据作为战略核心[4-5],是各国获得或维持其科研优势的重要资源。在我国,科学大数据作为国家大数据战略基石[6],被视为"传播速度最快、影响面最宽、开发利用潜力最大"的"重要的战略资源"[7-8],其安全问题因此引起广泛关注,尤其是科学数据的产权、主权、共享安全问题。鉴于我国尚未就数据产权立法,本文试图以现行法中与科学数据产权存在大量交叉关系的知识产权为切入点,以法律规范分析为基础,运用法律解释方法,厘清知识产权与科学数据的关系,并结合国内外实践案例探索保障科学数据产权、主权与共享安全的产权管理具体方案。

1 科学数据安全问题研究综述

数据(Data),是对客观事件进行记录并可以鉴别的符号。在计算机科学中,数据是指所有能输入到计算机并被计算机程序处理的符号[9]。科学数据(Research Data 或 Scientific Data),也称科技数据[10],陈传夫[11]、王学勤[12]等学者,以及各个科研机构对科学数据有不同定义[13-19]。为避免歧义,本文所称科学数据,特指在科研活动中取得并用于科研活动的、对客观事件进行记录并可以鉴别的符号,包括原始数据及其衍生数据。

数据安全,在不同维度有不同的内涵。从体制角度来看,指通过制度来保障数据免遭遗失、泄露、非法获取、修改、利用、毁损;从形态角度来看,指数据存储

* 本文系国家科技图书文献中心(NSTL)自主项目"科学数据安全边界研究"(项目编号:2020XM44-2)和国家社会科学基金青年项目"图书馆数字化服务的著作权问题及其解决方案研究"(项目编号:17CTQ037)研究成果之一。
作者简介:罗娇(ORCID:0000-0003-4949-5211),副教授,博士;刘细文(ORCID:0000-0003-0820-3622),研究员,博士生导师,通讯作者,E-mail:liuxw@mail.las.ac.cn。
收稿日期:2020-12-01 修回日期:2021-03-18 本文起止页码:38-46 本文责任编辑:杜杏叶

罗娇，刘细文. 知识产权视角下科学数据安全管理的策略选择[J]. 图书情报工作,2021,65(12):38－46.

和传输安全;从内容角度来看,指硬件、软件、系统、内容安全;从主体角度来看,指国家、社会、企业和个人数据安全[20];从操作角度来看,指数据完整性和数据利用的安全[21]。从国外立法实践来看,一般认为保持数据的机密性、完整性和可用性是数据安全必须满足的三个基本要素[22];从我国国家层面的立法来看,尚未对数据安全进行定义,也未对科学数据安全进行定义。

我国学者以科学数据的机密性、完整性、可用性为基础,从不同视角探讨了科学数据的安全问题,并提出安全保障措施。李善青等从技术视角,提出以完整性校验技术、访问控制技术、数据加密技术、隐私保护技术、安全审计技术综合加强科学数据安全[23]。杨燕等从行为的视角,通过构建科研人员在科学数据生成、采集、存储、描述、分析、利用、汇交等过程中的行为规范来保障数据安全[24]。肖冬梅等从主体的视角,结合云环境,提出以发展云安全产业、构建数据保护法律体系和政府实施数据监管,应对科学数据安全风险[25]。盛小平等从治理的视角,指出要采取加强立法、制定政策、构建标准、创建系统、借助技术、建设体制机制等多种措施进行科学数据安全治理[22]。刘桂锋等从管理的视角,强调可以基于科学数据生命周期,从"制度层""基础设施层""数据素养层"和"实施层"构建科学数据安全框架[26]。以上研究从不同的视角解读科学数据的安全问题,就保障科学数据安全提出了深刻见解,对促进我国科学数据安全发展具有重大价值。但相关研究多关注探讨各种措施,而缺少分析每种措施的法律依据和权利基础。即使涉及法律分析,其阐释也侧重于应然层面(法律应该是什么样的),而较少探讨实然层面(法律实际是什么样的)。

国外的研究与实践较重视从产权的角度维护科学数据安全。相关研究阐释了知识产权与合同对数据安全共享的重要价值[27],并对保障科研人员的权益、承认科研人员的贡献发挥了重要作用[28]。哈佛大学以知识产权框架下的科学数据产权政策作为保障其科学数据安全的重要支撑[29];康奈尔大学强调从版权的角度管理数据产权[30];芬兰科学数据中心借助版权和合同对科学数据进行产权管理[31];堪称数据开放典范的美国国立卫生研究院(National Institutes of Health,简称NIH),也在最新的数据共享政策中强烈鼓励科研人员使用NIH建立的数据库来保存和共享科学数据[32],以实现NIH对数据产权的事实控制。欧洲分子生物学实验室公共数据库(European Molecular Biology Laboratory-European Bioinformatics Institute,简称EMBL-EBI)甚至在其使用条款中直接警示用户重视数据产权问题,尤其需重视数据产权法律上的不确定性对科学数据重用的干扰[33]。

鉴于具备数据掌控能力和分析能力是科学数据安全的重要方面,具备数据掌控能力和分析能力又以实际控制科学数据为前提,而实际控制科学数据的直接法律依据就是对数据享有产权,本文尝试以现行法中距离科学数据最近的权利——知识产权作为切入点,分析科学数据的产权、主权和共享问题,以期填补科学数据安全管理在此方面的研究空白。

2 科学数据产权管理面临的挑战

2.1 科学数据的产权缺位问题

科学数据归谁所有、由谁使用、使谁获益,关涉科学数据的产权制度,因产权制度缺位带来的产权不清晰、产权流失,是我国科学数据安全管理的核心问题。

2.1.1 产权缺位致使产权不清晰

我国数据产权立法缺位,使数据产权的归属与分配无法可依。我国《民法典》第127条将数据纳入保护范围,但该条仅是数据保护的引致规范,数据产权问题处于立法空白。2018年国务院办公厅印发《科学数据管理办法》,明确了我国科学数据管理的原则、职责,但并未涉及产权制度,科学数据产权缺乏顶层设计。司法实践中,我国法院多以《合同法》和《反不正当竞争法》来解决数据产权问题,但是两种路径均有局限。就《合同法》而言,一方面,合同救济的前提是合同既已订立,而科学数据采集、加工、共享与利用多依赖于机构政策、行业指南或学术伦理进行,通常缺乏全面、严密的合同安排;另一方面,鉴于合同的相对性,合同一般不约束合同之外的第三人,仅有合同不足以应对第三人对科学数据的侵害问题。就《反不正当竞争法》而言,一方面,该法仅适用于市场竞争主体,而科学数据不一定与市场竞争相关;另一方面,该法并未将数据产权纠纷定性为一项不正当竞争行为,法院多以《反不正当竞争法》第二条"一般性条款"来处理数据产权问题,但是,一般性条款缺乏具体的适用要件和损害后果,以此处理数据产权问题充满不确定性。

2.1.2 产权不清晰致使产权流失

科学数据产权流失主要体现在三个方面:一是科学数据未进行妥善保存;二是科学数据被不当披露;三是科学数据被收编入他人权利的控制范围。首先,产权制度缺位难以通过排他性权利的形式对科学数据的享有、流通和使用形成强有力控制,因此妥善保存数据

便是事实上掌控数据的直接途径,也是后续开发、利用、交换、交易数据的基础条件。未得到妥善保存的科学数据,意味着失去事实上的控制力,数据会面临流失问题。早在 2008 年,我国科研行政管理部门便提出要对国家项目中形成的科学数据强制汇交,这意味着彼时相关部门已经意识到投入了大量国家公共资金的科研项目所产生的科学数据未得到妥善保存而毁损遗失或流向国外[8]。其次,产权制度缺位,使一些重要数据不得不以商业秘密的形式进行保护。以商业秘密的形式保护科学数据,数据必须具备秘密性、保密性、价值性。一旦数据被披露,即丧失秘密性,不再予以保护,这直接阻断了科学数据随论文发表、科学数据本身作为数据论文发表、科学数据开放共享等科学交流。再次,产权制度缺位,使科学数据一旦被收编入他人权利范围内,就会成为他人的权利。例如,科学数据被汇编入他人享有著作权的数据库中,会成为他人数据库作品的一部分。又如,数据被纳入他人专利技术方案中,会成为他人专利权的内容。一旦出现上述情形,再利用这些科学数据则会受制于他人权利。

2.2 科学数据的主权缺失问题

科学数据的主权,体现为国家对科学数据的掌控能力与管辖权力。国家主张主权,以本国国民实际占有数据或享有数据产权为前提,对于既没有实际占有数据,又没有纳入产权控制的数据,国家对此主张主权缺乏法律依据。因此,数据主权是数据产权得以实现的基础和前提,数据产权又为数据主权得以维护提供了有力依据[34]。在我国,科学数据的主权问题体现在由于产权缺位所带来的主权缺失上,具体表现为科学数据向国外流失。

科学数据外流问题引起了多位学者的关注[35-36]。现有研究显示,科学数据外流问题主要体现在三个方面,一是国外学术期刊、数据期刊、数据平台、数据中心的“吸虹”效应明显,二是国际科研合作中数据跨境传输不合规不合法,三是云存储与云计算使数据的国界越发模糊,对数据主权提出挑战。例如,国外一些顶级学术期刊的投稿政策要求作者将支撑其论文的数据提交到指定数据库中(通常是国外数据库),科研人员为了发表论文不得不遵守此类政策[37]。又如,一批国外高影响因子的数据期刊,利用其先发优势吸引了我国大批学者进行科学数据发表或存储到其指定平台上[37]。再如,一些国外知名的科学数据平台或科学数据中心也正在吸引着我国科学家将其数据存储到这些平台[37]。以我国的生命健康大数据为例,我国学者发表了大量的 SCI 论文,其中绝大部分的数据只能被递交到 NCBI、EBI 等国际知名数据库,NCBI 数据库中 25% 以上的数据来自我国,由于我国在科学数据管理上缺乏顶层设计,此类科学数据严重流失[38]。最后,如果科学数据未在国内指定机构汇交保存,随着云技术的广泛应用,用户可以通过互联网存储、访问云端数据,但真正掌握数据控制权的却是云存储服务提供商,一旦云服务因各种原因终止,则数据也可能因此流失。

科学数据大量外流所带来的安全问题不容小觑。一是科学数据大量外流会影响科学数据本身的安全问题。科学数据大量外流会导致我国科学家再利用此类数据时面临障碍,例如受制于他人的知识产权(如著作权、数据库权、专利权等),或者受制于他人的政策条件(如需要经过复杂漫长的伦理审查,被迫接受不合理的、过于严苛的使用条件等)。二是科学数据大量外流会影响科学数据利用的安全问题。科学技术知识在各种生产要素中占据显著地位,其毁灭性力量如果被极端分子不当运用,会带来极大风险[39]。例如利用基因数据、技术扩大某一族群的基因缺陷[40]。

2.3 科学数据的共享受阻问题

科研不是孤岛,数据活于流动,数据共享是科学事业的基础[41]。“科学数据资源是国家重要的财富,也是生产力,必须把科学数据流动起来,共享应用”[8]。

早在 2001 年,我国科技部便提出“实施科学数据共享工程, 增强国家科技创新能力”的建议, 随后各领域的科学数据共享工程陆续启动[42]。时至今日,我国科学数据共享取得了令人瞩目的成就,但进一步发展仍受制于文化、机制和技术等根源性问题[43]。在国外,即使被认为是科学数据共享成功典范的 NIH,也在其最新的数据共享政策中承认,法律、道德、技术因素对保存和共享数据的能力形成制约[32]。

产权缺位是科学数据共享的主要阻碍因素之一,如知识产权难以界定[44]、知识产权专有性和科学数据共享性存在冲突[45]等。本文认为,产权制度缺位阻碍科学数据共享的原因有三:一是科学数据持有者担心开放共享将丧失对访问数据的控制[22],失去科研竞争优势;二是数据重用的有效性受到干扰,因为数据权属存在法律上的不确定性,重用容易引起纠纷;三是科学数据的贡献者得不到应有的贡献承认,进而失去提升影响力、获得科研资助或科研奖励的潜在机会。结果是,我国数据库资源总数位居世界前列,但共享利用效率极低,缺少处于国际领先地位的大规模优质数据中心;引用数超过 500 次的数据库凤毛麟角,超过 1 000 次的更是为零[38]。

应该共享的数据没有共享,会形成资源浪费;不应该共享的数据被违规共享,就会影响数据安全。2015年华大基因公司、复旦华山医院在与英国牛津大学开展的国际合作研究中违规共享我国基因数据[46－47],2016年苏州药明康德公司违规与外方共享我国血清数据[48],不仅造成了数据流失,也敲响了数据共享安全的警钟。因此有学者指出,"科学没有国界,但科学数据是有国界的,平衡、适度的开放,维护科学数据开放的合法界限,建立数据共享和对外交流的安全审查机制也是科学数据开放共享的一个重要问题"[49]。可见,共享科学数据虽是趋势,但并不意味着放弃数据的产权和主权,明晰产权、主张主权,对保障科学数据安全共享具有重要意义。

3 将知识产权作为保障科学数据产权安全的核心手段

3.1 利用知识产权界定科学数据的产权

保障科学数据安全的关键是提升对科学数据的实际控制力,明晰产权、主张主权、阻止违规共享将有助于增强科学数据控制、开发、传播的主动权和主导权,获得大国科研博弈的竞争优势。在当前数据产权制度缺位的情况下,以知识产权为基础来界定科学数据的产权边界是可行举措。就我国当下法律环境而言,具体有著作权、专利权、商业秘密三种类型的知识产权可以作为科学数据保护的权利基础。

3.1.1 以著作权作为科学数据产权界定的基础

著作权是赋予科学数据产权的重要权利类型。就我国著作权法而言,可以从三个层面为科学数据提供产权保护:一是可以对数据赋予独创性表达来获得著作权保护。例如图纸或照片,绘制者和拍摄者可以通过构图、光影、聚焦方式来体现独创性,使相应的图纸或照片获得著作权。否则,原始的观察和实验数据属于"事实",不受著作权保护。二是可以对数据进行独创性的选择和排列而获得著作权。例如,如果科研人员在选择字段名称和排列其顺序时体现了独创性,那么即使是以 Excel 电子表格形式体现的数据也将获得汇编作品著作权,但著作权仅限于数据集层面,并不延及数据内容本身。这意味着如果另一位科学研究人员以别的字段名称重新排列电子表格中的这些数据,并不侵犯原数据集的著作权。三是对数据的阐释、数据可视化、元数据等产物具有足够的独创性可以获得著作权。创建科学数据的可视化、图形、图表、图片或以其他形式"处理"得到的产物,通常涉及处理者的独创性表达而获得著作权保护,如"黑洞照片"。

3.1.2 以商业秘密作为科学数据产权界定的基础

对于属于秘密信息的科学数据,可以作为商业秘密进行保护。根据我国《反不正当竞争法》的规定,商业秘密需具备秘密性、保密性和价值性。因此科学数据的创建者或控制者必须采取适当的保护措施,使数据处于秘密状态,才能以商业秘密的形式保护数据。实践中,通常通过物理控制(如设置门禁等)或技术控制(如进行加密或设置密码等)来保护科学数据的秘密状态。此外,数据的秘密性与保密性也可通过保密协议得到保证。需要注意的是,以商业秘密形式来保护科学数据与科学数据开放共享相冲突,开放共享意味着丧失商业秘密权利。

3.1.3 以专利权作为科学数据产权界定的基础

当科学数据具有实用性时,可能构成专利的一部分或本身可申请专利,如基因数据或与其相关的专利。以专利权保护科学数据有两种方式:一是申请专利。将科学数据纳入专利的保护范围内。对于此类本身可构成专利或专利的一部分的数据,需确保数据不被公开。在获得专利之前应制定数据保密协议,并严格按照保密协议向特定对象披露数据,以确保数据公开而成为现有技术;二是阻止他人申请专利。一些科研人员希望其数据不受包括专利在内的法律限制,使数据能够持续共享和重用。对此,将数据发布到公共领域就足以使相关数据成为现有技术,从而阻止他人成功申请专利。但是,即使在公开发布数据的情况下,他人也有可能使用所披露的数据作为其申请专利的技术方案的一部分。如果该发明被授予专利权,则随后的专利权可能足以阻止构成该发明的这部分科学数据的实际使用[50－52]。实践中,一般采用合同的方式来使科学数据免受该类专利权的限制。例如通过在线数据库发布数据会要求访问数据库的用户签订"数据使用的点击同意协议",该协议可禁止基于某些数据进行专利申请,也可以允许进行专利申请,但前提是该专利不得具有限制性,并且必须允许对该数据进一步使用。又如,约定利用科学数据获得专利的发明或设计,对于其中所包含的科学数据,根据"CC 协议"等开放许可协议许可他人自由使用[53－54]。

3.2 利用合同完善科学数据的产权安排

虽然通过一定形式可以将科学数据纳入著作权、专利权和商业秘密等知识产权的保护范围内,但是知识产权和科学数据安全是交叉关系而非重叠关系,知

识产权并不能覆盖所有的科学数据,科学数据的产权边界需要借助合同才能进一步明确,见图 1。保密协议、许可协议、访问协议是保护科学数据安全的基本合同类型。

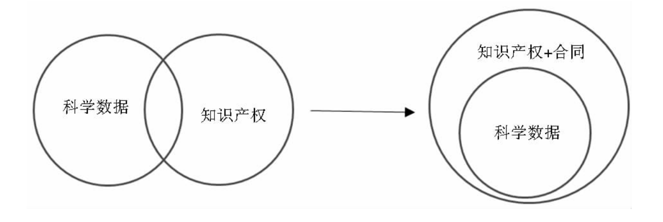


图 1 知识产权、合同与科学数据产权的关系

保密协议通过约定保密义务和保密措施,以控制信息被公开的方式来保护机密数据,从而使相关数据保持秘密状态。保密协议通常需约定有关机密信息的权利人、被视为机密的信息、信息被允许的使用范围、对披露信息者施加的义务以及约定不遵守保密义务的后果。

访问协议通常在科研人员或科研组织可以控制其数据存储的地方使用。访问协议通常包含以下内容:拟被访问的数据范围;被允许访问数据的人员范围;访问权是否能转让给第三方(一般数据方会声明不允许转让访问权);被访问的数据是否限制使用目的(如约定数据不能用于商业目的,或约定商业使用的须向数据创建者补偿一定比例的费用或利润等);免责声明(数据方对数据中的任何错误不负责任),以及访问方不遵守访问协议的后果。此外,访问协议还有一大作用——当数据丧失秘密性、无法用保密协议来控制数据的访问与使用时,可以利用访问协议来控制他人对数据的访问和使用。

许可协议授予他人处理数据库或数据集的权限,是科学数据安全共享的关键环节。如果没有授权许可,数据库的管理者将无法合法复制并向公众提供任何受专有权利保护的科学数据或数据集,用户也无法使用或以其他方式处理数据。实现科学数据的安全共享,至少需要两份授权许可协议:一是数据权利人对平台(多为数据库)存储、提供数据的授权许可;二是数据权利人对最终用户使用数据的授权许可。许可协议可以在许可期限、被许可的权利、许可的地域范围、行使权利的目的等方面约定限制条件。

4 管理与政策建议

4.1 加强科学数据产权管理

以知识产权为基础,借助合同来保障科学数据的产权安全,可通过设计科学数据权利归属条款、许可授权条款、保密条款和利益分配条款予以实现。

4.1.1 在权利归属方面,区分基础数据和衍生数据

对于基础数据,应明确谁是基础数据内容、数据库及其包含的一切算法的所有者。如果不提前声明这一点,将留下一个明显的合同漏洞——相关数据的归属可能受到挑战,如对方主张数据库不受法律保护等。对基础数据二次利用产生的衍生数据,如果不明确其归属,不仅会影响基础数据的拥有者将数据授权给其他主体,还会让数据重用者在数据开发上的投入打水漂。因此可在数据重用协议中,协商衍生数据归属于许可方或被许可方或由双方共有。无论如何约定归属,均要避免将自己拥有产权的数据与获得授权的数据进行合并,否则可能在无意间将自己的数据让渡给对方。

4.1.2 在授权使用方面,明确授权的类型、范围、限制条件

对于授权方,应考虑以何种方式对数据进行授权,如独占许可、排他许可、普通许可等。在某些情况下,被授权方的技术不能实现对数据的开发利用,被授权方会希望向第三方发放分许可,让第三方协助自己进行数据开发利用。对此,授权方可对被授权方的分许可约定限制,如约定分许可的行为仅出于被授权方利益且仅代表被授权方。此外,由于不同数据集有不同的使用限制,数据合并会带来新的风险,因此授权方还应当考虑是否允许被授权方将获得许可的数据与其他任何未获得授权的数据进行合并。最后,为了避免遗漏,授权方还可约定对相关数据的许可仅限于协议列出的授权范围,协议未注明的一切权利归授权方所有。对于被授权方,首先应当明确授权条款是否能够、如何能够让自己用相关数据实现目前设想的以及将来可预见的目标。如果被授权方获得的是普通许可,则需考虑是否要求授权方不得将数据再许可给被授权方的竞争对手,因为一旦竞争对手掌握了同样的数据,则被授权方的竞争力将受到影响。

4.1.3 在涉密数据方面,设置数据访问与重用协议中的保密条款

妥善的保密条款至少应包含以下内容:一是要求各方承认数据的秘密性,除非获得明确授权,否则不得公开;二是要求各方对数据采取保密措施,如部署相应的物理、管理和技术保护措施等,以维持相应数据的秘密性;三是限制被许可方发放分许可的能力,如要求被许可方只能向第三方提供自己有权披露的数据,并要

求第三方负有同许可方与被许可方之间保密协议的相同保密义务。

4.1.4 在利益分配方面,平衡数据拥有者与数据重用者的利益

科学数据本身和科学数据的开发利用都有重大价值,都是资源密集型工作,需要大量的投资和资源。因此,许可费用的分配机制对维系平衡至关重要。一方面,许可费用不能过高,这样才会不断有新的被许可方加入数据开发的生态系统中。另一方面许可费用也不能过低,这样才能确保许可方在数据上的投资得到回报。利益分配的设计应当维持一种平衡,既不过分提高数据开发利用的进入成本,导致相关研发停滞,也不过分降低数据拥有者的获利预期,使数据方不愿意继续收集、维护和提供数据。具体可通过创新许可费支付模式,如建立预付模式、分期付款模式、设置分许可利润分享条款等。此外,数据开发利用不仅需要大量的投资,也需要大量的时间。一旦开发者对数据产生了依赖,即形成"卖方市场",开发者以合理成本获取数据的可能性将大打折扣。因此有必要在许可费用条款的设计中,加入价格保护条款,如要求数据拥有方在一个固定期限内不得提高许可费用;若固定期限过后提高费用,需留出足够的时间书面通知对方,且费用涨幅不得超过一定比例。

4.2 增强科学数据主权意识

4.2.1 在国家层面构建科学数据管理的顶层设计,重视数据产权安全

从国家层面制定科学数据产权管理的国家原则是国外科学数据管理工作的一个有效经验。例如澳大利亚研究理事会(Australian Research Council,简称 ARC)联合澳大利亚政府科学、工业和资源部、澳大利亚知识产权局制定的《澳大利亚国家公共研究的知识产权管理原则》强调实现最大化国家利益和公共研究的投资回报[55]。又如,经合组织科学技术政策委员会《关于获取公共资金资助的研究数据的声明》强调,科学数据访问机制应充分注意国家安全、隐私和商业秘密的国内法要求,明确参与数据活动的各方主体责任[56]。

4.2.2 在机构层面建立科学数据管理与共享政策,避免数据违规外流

机构可视情况根据数据类别建立科学数据管理与共享政策,如世界卫生组织(World Health Organization,简称 WHO)分别制定了"突发公共卫生事件下数据共享政策"[57]和"非突发公共卫生事件下数据共享政策"[58]。同时,机构应该在其数据管理与政策中规定科学数据的本地化存储等安全管理措施。例如美国国立卫生研究院在其最新的数据管理政策中,强烈鼓励尽可能使用其已建立的存储库来保存和共享科学数据,并要求科研人员在数据管理计划中说明如何管理研究项目生成的科学数据,以及共享哪些科学数据和随附的元数据[59]。

4.3 促进科学数据开放共享

4.3.1 利用许可协议实现科学数据的安全共享

许可协议可以适用于两种情况,一是当科学数据之上覆盖有一项或多项知识产权时,权利人可以通过许可协议授权他人重用数据。二是科学数据虽然未纳入知识产权的保护范围,例如按时间顺序排列的大量传感器数据集合,但获取这些数据需要从与"访问协议""使用条款"相关联的数据平台进行访问或下载,此时,这些"访问协议""使用条款"仍可作为"合同"来对数据的获取与重用作出限制。第二种情况相当于对处于公有领域的数据施加了合同约束,因此,此类"访问协议""使用条款"的设计应以科学数据开放为原则,不开放为例外。

4.3.2 利用法律机制清除科学数据开放共享的权利障碍

对于商业秘密权,通过公开披露便可清除数据上的商业秘密权。其他形式的共享,只能在特定的保密形式或保密协议下,在约定的范围内共享数据。对于专利权,如上所述,公开披露可使数据丧失新颖性,从而限制或破坏与数据相关的发明的可专利性。当专利涵盖科学数据时,相应数据虽然随专利申请文件公开,但对数据的重用若构成实施相关专利,则需获得专利权人的许可。对于著作权,在开放数据实践中,一般使用"CC 协议"等开放许可协议,对受著作权保护的数据或数据集设置访问和使用的权限和条件。"CC 协议"的实质是授予用户从事受著作权控制的使用行为,且不对用户附加其他义务,从而实现科学数据的开放共享。开放知识基金会(Open Knowledge Foundation,简称 OKFN)将此类协议的好处描述为"允许他人自由广泛地许可作品;他人每次希望使用或传播作品时,不用再费时地寻求许可;鼓励他人不断为作品增值;鼓励他人根据原始作品或从原始作品中创作出新作品"[60]。除了 CC 协议外,GNU 自由文档许可(GFDL)、开放数据许可协议(ODC)[61]也被广泛运用与开放数据实践中。大多数开放许可协议都支持在线访问,并可以轻松地附加到数据上。

特定领域科学数据的开放共享,例如医学领域,还

需关注隐私和个人信息问题。对于包含个人信息、隐私等敏感信息的数据，公开和重用可能受到隐私权和个人信息权的限制[62]。最佳方法是在可行的情况下采用最高等级的隐私合规标准，全程监管数据收集和共享，以保证其安全性和私密性，并且共享前可采取授权许可措施或脱敏措施，例如寻求数据所涉当事人的同意，对数据进行匿名化或模糊化处理等。一旦发现脱敏处理失效或可能失效，则立即撤回对相关数据的使用。此外，随着数据分析技术的精进，传统的匿名化或模糊化措施可能失效，建议对此类数据只开放数据分析结果，不开放或谨慎开放原始数据。

5 结语

数据时代，科学研究的第四范式——数据密集型科学研究应运而生，科学数据成为国家的重要战略资源，维系其安全势在必行。科学数据安全的核心在于提升对科学数据的掌控能力与应用能力，以保障科学数据的安全存储、安全使用、安全共享和安全服务为前提。然而，长期以来的产权缺位、主权缺失和共享缺失，埋下了科学数据安全隐患。

在当前数据产权立法尚处空白的情况下，知识产权作为距离数据权益最近的权利，以知识产权为基础，借助合同来明确科学数据的产权、维护科学数据的主权和规范科学数据的共享，具有一定的可操作性。当然，知识产权本质上是私权，其以“理性经济人”为逻辑起点，以“权利本位”为价值底色，与数字时代“去中心、扁平化、无边界”的内在机理和“开放、合作、共享”的科学精神相去甚远，以知识产权维系科学数据的安全边界只是权宜之计。“法与时转则治”，开创“数据权”，进行全新的制度构建，也许才能实现“数尽其用、定纷止争”，捍卫科学数据安全。

参考文献：

[1] 邓仲华，李志芳. 科学研究范式的演化——大数据时代的科学研究第四范式[J]. 情报资料工作，2013(4)：19－23.
[2] HEY T, TANSLEY S, TOLLE K. The fourth paradigm: data-intensive scientific discovery [M]. Redmond: Microsoft Research, 2009.
[3] 郭华东，王力哲，陈方，等. 科学大数据与数字地球[J]. 科学通报，2014，59(12)：1047－1054.
[4] CUKIER K, SCHONBERGER M V. The rise of big data: how it's changing the way we think about the world[J]. Foreign affairs, 2013(92)：28－33.
[5] OBAMA Whitehouse. Big data research and development initiative [EB/OL]. [2020－11－01]. https://obamawhitehouse.archives.gov/blog/2012/03/29/big-data-big-deal.
[6] 郭华东. 科学大数据——国家大数据战略的基石[J]. 中国科学院院刊，2018，33(8)：768－773.
[7] 中国政府网. 保障安全、突出共享、支撑创新——科学数据，如何科学管理[EB/OL]. [2020－11－01]. http://www.gov.cn/zhengce/2018-04/08/content_5280429.htm.
[8] 中国政府网. 科学数据是重要战略资源[EB/OL]. [2020－11－01]. http://www.gov.cn/zhengce/2018-04/06/content_5280211.htm.
[9] 叶斌，黄洪桥，余阳. 信息技术基础[M]. 重庆：重庆大学出版社，2017.
[10] 孙九林，黄鼎成，李晓波. 我国科技数据管理和共享服务的新进展[J]. 世界科技研究与发展，2002，24(5)：15－19.
[11] 陈传夫. 中国科学数据公共获取机制：特点、障碍与优化的建议[J]. 中国软科学，2004(2)：8－13.
[12] 王学勤，STOUT A, SILVER H. 建立数据驱动的 e-Science 图书馆服务：机遇和挑战[J]. 图书情报工作，2011，55(3)：80－83.
[13] NSF. NSB-05-40 Long-lived digital data collections: enabling research and educating in the 21 century [EB/OL]. [2020－11－01]. http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf.
[14] ANU. ANU data management manual: managing digital research data at the Australian National University [EB/OL]. [2020－11－01]. http://regnet.anu.edu.au/sites/default/files/Iiles/ANU_Data_Management_Manual.pdf.
[15] OECD. OECD principles and guidelines for access to research data from public funding [EB/OL]. [2020－11－01]. http://www.oecd.org/dataoecd/9/61/38500813.pdf.
[16] University of Cambridge. Explanation of terms [EB/OL]. [2020－11－01]. http://www.lib.cam.ac.uk/preservation/incremental/glossary.html.
[17] 农业科学数据共享中心. 简介[EB/OL]. [2020－11－01]. http://www.agridata.cn/homepage/chintro.asp.
[18] 测绘科学数据共享服务网. 简介[EB/OL]. [2020－11－01]. http://sms.webmap.cn/.
[19] 科学数据共享概念与术语[EB/OL]. [2020－11－11]. http://www.sciencedata.cn/pdf/2.pdf.
[20] 齐爱民，盘佳. 数据权、数据主权的确立与大数据保护的基本原则[J]. 苏州大学学报(哲学社会科学版)，2015，36(1)：64－70，191.
[21] 张尼，张云勇，胡坤. 大数据安全技术与应用[M]. 北京：人民邮电出版社，2014.
[22] 盛小平，郭道胜. 科学数据开放共享中的数据安全治理研究[J]. 图书情报工作，2020，64(22)：25－36.
[23] 李善青，郑彦宁，邢晓昭，等. 科学数据共享的安全管理问题研究[J]. 中国科技资源导刊，2019，51(3)：11－17.
[24] 杨燕，阮建海. 基于科研过程的科学数据安全行为研究[J]. 知识管理论坛，2019，4(4)：218－231.
[25] 肖冬梅，孙蕾. 云环境中科学数据的安全风险及其治理对策[J/OL]. 图书馆论坛，2021(2)：1－10.

[26] 刘桂锋,阮冰颖,包翔.数据生命周期视角下高校科学数据安全内容框架构建[J].情报杂志,2021,40(2):1-8.
[27] CARROLL M W. Sharing research data and intellectual property law: a primer [EB/OL]. [2020-03-12]. https://doi.org/10.1371/journal.pbio.1002235.
[28] CONTRERAS, J L. Data sharing, latency variables and science commons[J]. Berkeley technology law journal, 2010,25(11):1601-1672.
[29] Harvard University. Research data ownership policy [EB/OL]. [2020-03-12]. https://vpr.harvard.edu/data-ownership.
[30] Cornell University. Research data management service group data management planning best practices [EB/OL]. [2020-03-12]. https://data.research.cornell.edu/content/intellectual-propert.
[31] FSS. Finnish social science data archive data management guidelines [EB/OL]. [2020-03-12]. https://www.fsd.tuni.fi/en/services/data-management-guidelines/.
[32] NIH. Final NIH policy for data management and sharing [EB/OL]. [2021-03-12]. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html
[33] EMBL-EBI. Terms of use of the EBI Services [EB/OL]. [2020-03-12]. http://www.ebi.ac.uk/about/terms-of-use.
[34] TRACHTMAN P J. Cyberspace, sovereignty, jurisdiction and modernism[J]. Indiana journals of global legal studies, 1998,5(2):561-582.
[35]曾建勋,杨代庆.关于扭转我国科技论文外流局面的政策性思考[J].编辑学报,2020,32(6):600-604.
[36] 李洋,温亮明.我国科学数据外流:表现、问题与对策[J].图书馆杂志,2019,38(12):72-81,115.
[37] 国家科技基础条件平台中心.国家科学数据资源发展报告(2016)[M].北京:科学技术文献出版社,2016.
[38] 鲍一明,薛勇彪.生命与健康大数据现状和展望[J].中国科学院院刊,2018,33(8): 861-865.
[39] 钟灿涛.开放与保密:科技信息传播控制及其对创新的影响——以美国科技信息传播控制机制为例[J].科学学研究,2013,31(3):335-343.
[40] VOGEL G. German law allows use of DNA to predict suspects' looks[J]. Science, 2018, 360(6391): 841-842.
[41] RAI K A. Regulating scientific research: intellectual property rights and the norms of science[J]. Northwestern university law review, 1999,94(1):77-152.
[42]杨兰.开放科学数据 实现共同发展[EB/OL]. [2020-11-01]. http://finance.china.com.cn/roll/20140424/2358833.shtml.
[43] PENG C, SONG X, JIANG H, et al. Towards a paradigm for open and free sharing of scientific data on global change science in China [EB/OL]. [2021-02-11]. https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ehs2.1225.
[44] 刘润达,孙九林,廖顺宝.科学数据共享中数据授权问题初探[J].情报杂志,2010,29(12):15-18.
[45] 徐先东,朱雪忠.知识产权专有性与科学数据共享性的冲突研究[J].武汉理工大学学报,2007(7):164-167.
[46] 中华人民共和国科学技术部.行政处罚决定书:国科罚〔2015〕1号. [EB/OL]. [2020-11-01]. http://www.most.gov.cn/bszn/new/rlyc/xzcf/201810/ t20181011_142042.htm.
[47] 中华人民共和国科学技术部. 行政处罚决定书:国科罚〔2015〕2号. [EB/OL]. [2020-11-01]. http://www.most.gov.cn/bszn/new/rlyc/xzcf/201810/ t20181011_142043.htm.
[48] 中华人民共和国科学技术部. 行政处罚决定书:国科罚〔2016〕1号. [EB/OL]. [2020-11-01]. http://www.most.gov.cn/bszn/new/rlyc/xzcf/201810/ t20181011_142045.htm.
[49]王瑞丹,杨静,高孟绪.加强和规范我国科学数据管理的思考[J].中国科技资源导刊,2018,50(2):1-5.
[50] GITTER M D. Resolving the open source paradox in biotechnology: a proposal for a revised open source policy for publicly funded genomic databases [EB/OL]. [2020-11-01]. http://ssrn.com/abstract=901994.
[51] EISENBERG S R, RAY K A. Harnessing and sharing the benefits of state-sponsored research: IP rights and data sharing in California's stem cell initiative [J]. Berkeley technology law journal, 2006,21(11):1187-1213.
[52] DRISCOLL C T. NIH data and resource sharing, data release and intellectual property policies for genomics community resource projects[J]. Expert opinion on therapeutic patents, 2005,15(1):1-8.
[53] CAMBIA. About biOS (biological open source) licenses and MTAs [EB/OL]. [2020-11-01]. http://www.bios.net/daisy/bios/licenses/398.html.
[54] JEFFERSON R. Science as social enterprise: the CAMBIA biOS initiative[J]. Innovations: technology, governance, globalization, 2006,1(4):13-44.
[55] The Australian research council et al. IP Australia, the national health and medical research council, national principles of intellectual property management for publicly funded research [EB/OL]. [2020-11-01]. http://www.arc.gov.au/pdf/01_01.pdf.
[56] OECD. Declaration on access to research data from public funding [EB/OL]. [2020-11-01]. https://legalinstruments.oecd.org/public/doc/157/157.en.pdf.
[57] WHO. Policy statement on data sharing by the world health organization in the context of public health emergencies [EB/OL]. [2021-02-10]. https://www.who.int/about/who-we-are/publishing-policies/data-policy.
[58] WHO. Policy statement developing global norms for sharing data and results during public health emergencies (as of 13 April 2016) [EB/OL]. [2021-02-10]. http://www.who.int/wer/2016/wer9118/en/ or www.who.int/ihr/procedures/SPG_data_sharing.pdf
[59] NIH. Policy for data management and sharing[EB/OL]. [2021-

02 – 10]. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html.
[60] OKFN. Open knowledge foundation [EB/OL]. [2020 – 11 – 01]. http://www.opendefinition.org/guide? action = print.
[61] 罗娇. 开放数据的著作权解决方案——ODC 协议研究[J]. 图书情报工作,2017,61(9):23 – 32.
[62] 罗娇. 大数据环境下个人信息保护法律问题研究[J]. 图书馆,2018(5):31 – 36.

作者贡献说明:
罗娇:设计论文框架,撰写论文初稿;
刘细文:提出研究命题,设计研究框架,修改文稿及定稿。

Strategies for Scientific Data Security Management from the Perspective of Intellectual Property

Luo Jiao[1] Liu Xiwen[2,3]
[1] College of Humanities and Development Studies, China Agricultural University, Beijing 100083
[2] National Science Library, Chinese Academy of Sciences, Beijing 100190
[3] Department of Library, Information and Archives Management, School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190

Abstract: [**Purpose/significance**] Scientific data security is a complex and dynamic concept that includes sovereignty security, property rights security, and sharing security, among which property rights security is an important prerequisite for sovereignty security and sharing security. In view of that China has not yet legislated on data property, it is of great significance to analyze the property rights relationship in scientific data security management by taking the intellectual property rights that have a cross relationship with scientific data property rights in the current law as the starting point. [**Method/process**] Aiming at the problem of scientific data's property, sovereignty, and sharing, using legal interpretation methods to analyze the relationship between intellectual property and scientific data, and to explore possible solutions under domestic and foreign practical cases. [**Result/conclusion**] For property, protect the property of scientific data by intellectual property rights and contracts with rights ownership clauses, authorization clauses, confidentiality clauses and benefit distribution clauses. For sovereignty, enhance the awareness of scientific data sovereignty by establishing the national top-level management and establish the institutional-level scientific data management and sharing policy to avoid data outflow illegally. For sharing, remove the right barriers to the open sharing of scientific data under the characteristics of different types of intellectual property rights, and realize the open sharing of scientific data through licensing agreements.

Keywords: scientific data intellectual property contracts data security